DOCUMENT RESUME

ED 276 747 TM 860 697

AUTHOR Goodison, Marlene
TITLE Pros and Cons of Paper and Pencil Tests for Teacher Assessment.
PUB DATE Jun 85
NOTE 8p.; Paper presented at the Annual Assessment and Policy Conference of the Education Commission of the States/Colorado Department of Education (16th, Boulder, CO, June 9-12, 1986).
PUB TYPE Viewpoints (120) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS College Entrance Examinations; Education Majors; Elementary Secondary Education; Evaluation Methods; Higher Education; *Licensing Examinations (Professions); Mastery Tests; *Teacher Certification; *Teacher Evaluation; Testing Problems; *Test Use; Test Validity
IDENTIFIERS Paper and Pencil Tests; *Teacher Competency Testing

ABSTRACT

With the increasing emphasis on teacher competency testing, it is important to examine the usefulness of the currently available paper and pencil tests. At least 40 states in the United States now have, or are planning to have, teacher tests for initial certification. Some states have tests for career advancement, and 20 mandate tests for admission to the teacher education curriculum. Such tests have both strengths and weaknesses. Test results cannot solve all of the problems of teacher selection, do not predict classroom performance, and must not take the place of informed decision making by experienced administrators. Good tests do, however, provide objective, standardized and reliable evidence that examinees possess the basic competencies judged necessary to perform competently as beginning teachers or education students. They assess those skills which are necessary for minimum competence, but not all skills which are sufficient for good teaching. Ideally, licensing tests should be used as pass-fail tests in conjunction with other criteria such as grades, experience, interviews, observation, and recommendations. Paper and pencil tests alone do not predict job performance; the use of simulated tasks and on-the-job evaluation add to predictive validity. (GDC)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED270747

Pros and Cons of Paper and Pencil Tests for Teacher Assessment

Marlene Goodison
Program Administrator
Teacher Programs and Services
Princeton, NJ

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.

☑ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

M. Goodison

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

PROS AND CONS OF PAPER AND PENCIL TESTS FOR TEACHER ASSESSMENT*

Marlene Goodison
Program Administrator
Teacher Programs and Services
Princeton, NJ

As recently as 1981 there were only 16 states that had mandated or were planning competency testing for initial certification of teachers. Today at least 40 states already have or are planning a mandatory testing program for initial certification, and a growing number have career ladder programs for teachers that include testing at professional levels other than preservice. Approximately twenty states mandate testing prior to admission to teacher training. It is obvious that the movement for testing is growing. What are its origins?

For some time parents have been expressing concern that their children are inadequately prepared for work or for higher education. Similarly employers complain that high school graduates cannot read instructions, do basic business math, address the public articulately or write a literate business letter. Legislators are frustrated at the cost of education, the complexity of its objectives, the lagging performance of students, and the lack of accountability in the educational system. The report of the President's Committee on Excellence in Education (1983) and other recent assessments of education in America, have served to underscore these complaints and citizens clamor for improvement.

For many, it appears that the quickest way to achieve the desired improvement in our children's education is to improve the teachers who impact it; and the most expedient way to do that is to test teachers for basic competencies before they can take their places in our classrooms. Across our country we see states responding to the public's demands and concerns by instituting mandatory competency testing programs for prospective teachers.

In the face of such strong demands for teacher testing -- a seemingly simple answer to a very complex problem -- it is critical that we keep in mind the strengths and weaknesses of the paper and pencil instruments of assessment that are available, and assure that use of such tests is appropriate. There are often unreasonable expectations of what such tests can do. The tests must not be made to carry a larger burden than they are capable of bearing; test results cannot solve all of the problems of teacher selection, do not predict performance in the classroom and must not take the place of informed decision-making by experienced administrators.

If we accept these limitations of what paper and pencil tests can do, can they be useful? Quite simply, the answer is yes. Good paper and pencil tests provide objective, standardized and reliable evidence, in the absence of the opportunity for direct observation, that examinees possess the basic competencies judged necessary to perform competently as beginning teachers or education students.

---

*Presentation at the 1986 Annual Assessment & Policy Conference of the Education Commission of the States in Boulder Colorado, June 1986.

In other words, they serve as testimony that the individual has attained a skill level in reading, writing, mathematics, pedagogy, or other subject areas that will permit him or her to operate competently in a designated role. They ensure that individuals have a firm grasp of the subjects tested.

There are three steps in a teacher's career when such information is most appropriate 1) to qualify entrants to teacher education programs, 2) to qualify job applicants, and 3) to qualify candidates for certification or licensing, including initial and renewal certificates and career ladder certificates.

In these applications, basic competency tests are advantageous and appropriate when used as necessary but not sufficient indicators of minimum competence. The phrase "necessary but not sufficient" becomes clearer to us if we consider an aspiring basketball player -- to qualify the player must at least master the techniques of dribbling the ball. If she can dribble the ball, she may be able to play the game well, provided she can also run fast, make baskets, pass accurately, guard the opposition, etc. But if she can't, at the least, dribble the ball, it doesn't matter if she can run fast, or do all the other things, she won't be able to play the game. Thus we can say, dribbling the ball is a necessary, but not sufficient requirement for play.

So it is with the basic competencies measured with paper and pencil teacher assessment tests, having them does not guarantee competency as a teacher. But, if one cannot read and write, one cannot begin to be a successful student in a teacher education program or a successful beginning teacher in the classroom. The "sufficient" qualities are another story -- and as we know, they require much more than existing pencil and paper tests to evaluate them.

How, then, should test scores be used. Essentially, they should be used as pass/fail instruments. They should always be used <u>with</u> other criteria to establish a pool of eligible candidates for certification, program entry, etc. Examinees -- whether prospective education students or applicants for certification -- who have attained the level of competence required by the examining agency on the tests, and other criteria such as GPA, past experience, etc., should then be evaluated on the basis of additional information, such as interviews, and recommendations.

In using tests as pass/fail instruments one must keep in mind the process of setting the pass/fail performance standard and its implications. It is imperative that these standards be set systematically and fairly since important decisions about individuals are based on them.

The score used to classify individuals as "qualified" or "not qualified" is an observed test score and as such there is measurement error that is associated with the test score(s). There will therefore be some classification error associated with the decision based on that score(s). There are two kinds of such classification errors -- false positives, passing someone who is not qualified, and false negatives, failing someone who is qualified. Policymakers must consider the consequences of misclassification and, depending on the situation, make every effort to either minimize the possibility of false positives or minimize the possibility of false negatives.(minimizing one, increases the likelihood of the other.) If one were selecting airline pilots where the consequences of a false positive would be disastrous, one would necessarily minimize the possibility of any false positives at the risk of creating false negatives, but that is not the case with teacher assessment. There are, and will be, other opportunities for evaluation and corrective action if there are any false positives. On the other hand, for those who may be false negatives they may have no further opportunity to show their competence. For just such reasons, several states have greatly reduced the chances of failing persons who are, in fact, qualified by setting their qualifying scores at one or two standard errors of measurement below an initially recommended qualifying score. In this way they adjust for the error in test scores.

If we know what the tests can do, what they are designed to measure, and how they should be used, we can address the issue of what tests cannot do, and how test scores should not be used.

No empirical relationship between basic competency test scores and actual job performance has been established. For this reason test scores should not be used for ranking examinees for selection. Test scores should never be used as the sole criterion in any decision process, nor for decisions regarding job retention, tenure, or dismissal. Educational Testing Service has stated unequivocally its position that putting an experienced fully certified teacher's professional career on the line solely on the basis of a one-time basic skills test is an injustice to the teacher and a misuse of the test. In no other occupation requiring licensure or state certification does such an ex post facto requirement occur.

Certainly, no one wants to keep an illiterate or otherwise incompetent teacher in the classroom, but in such cases, there are procedures for careful ongoing supervision and evaluation. And in the absence of improvement, termination, using due process, may be decided. A final decision must be based on all the relevant criteria--not only the necessary, but the sufficient too.

The state is empowered and obligated to protect the health, safety, and welfare of its citizens. No one questions the state's authority to require, that before persons are allowed to practice professions or

occupations, such as law, medicine, or dentistry, practitioners must demonstrate, on a test, that they have acquired the requisite knowledge to practice their profession. As with teacher competency testing, such medical exams, bar exams and other licensure exams do not predict how good a physician or attorney the examinee will be -- after all, all of the Watergate lawyers passed the bar exam and we all have read stories of incompetent physicians charged with gross malpractice yet no one would say that these malfeasances are reasons to abandon these licensing exams. Yet some say, that without predictive validity teacher competency tests should be abandoned. It is not reasonable to place a heavier burden on teacher tests than on other professional licensure tests. Licensure tests only assure us that the individuals who are licensed or certified have at least mastered the knowledge and skills necessary to perform competently in their chosen profession.

Certainly the school children of a state are entitled to at least as much consumer protection as its citizens receive in other professional areas. It is even more so the case when one remembers that a "free market" operates with other professions, that is, if you are dissatisfied with your physician or attorney, etc. you can take your business elsewhere. School children and their parents do not have this "Free Market" option in the public school system. States, therefore, have not only the right, but the obligation, to assure that school children within their jurisdictions have teachers who 1) have the basic knowledge of the subject they plan to teach; and 2) can demonstrate adequate skills in communications, literacy, and numeracy necessary to instruct children in an elementary or secondary classroom

To the citizens whose confidence in public education has faltered in the last decade, teacher competency testing offers a promising way to begin rebuilding public confidence, but we must be on guard that the superficial attractiveness of competency testing does not lead to misuse or misrepresentation of the tests.

If one considers the ways to measure job competence, there are four obvious approaches -- paper-and-pencil tests of necessary job knowledge and skills, job simulations, a set of actual job tasks to be performed by the candidate, and on-the-job evaluation. As one progresses from paper-and-pencil tests to on-the-job-evaluation, the predictive validity can be expected to increase, with on-the-job evaluation having the best predictive validity. However, just as we can not put pilots in a passenger airplane and wait to see if they can fly it to get the best predictive validity, we must not put unqualified persons in the classroom and then wait to see if they can teach our children. We can not use performance of job tasks because there is not yet a consensus in the profession on what might constitute a set of specific teacher's job tasks or how these tasks should be evaluated. We don't yet have adequate job simulations for teachers that can be used for summative evaluations. We do have valid, reliable, paper-and-pencil tests of necessary job knowledge.

If we remember that these paper and pencil teacher competency tests have their limitations and are not expected to predict success in the classroom, but to assure the public that the teachers in our classrooms possess <u>at a minimum</u> the necessary basic skills, then the paper-and-pencil tests will do the job for us they were meant to do. They can be the first step on the journey to educational improvement.

REFERENCES

Galambos, Eva C. "Testing Teachers for Certification and Recertification" National Commission on Excellence in Teacher Education (ED) Washington, D.C., 10/84.

Goodison, Marlene. "Teacher Competency Testing: What It can and Cannot Do," presented at the 1985 Annual Assessment and Policy Conference of the Education Commission of the States in Boulder Colorado, June 1985.

Goodison, Marlene. "Teacher Testing: Improving the Outcomes," presented at the Annual Meeting of the American Educational Research Association and the National Council of Measurement in Education in San Francisco, April, 1986.

Millman, Jason (1981, Fall) Protesting the Detesting of PRO Testing. Measurement in Education, 12,(3).

NTE Policy Council. Guidelines for Proper Use of NTE Tests. Princeton, NJ: Educational Testing Service, 1985

Rankin, Earl F., & Lacefield, Warren E. (1985) An Assessment of the Minimum Reading Comprehension of Prospective Teachers as Defined by a De Facto Standard. Research in Higher Education 22 (4), 381-394